Systems & Control Letters 23 (1994) 137–148
North-Holland

White noise theory of robust nonlinear filtering with correlated state and observation noises

Arunabha Bagchi

Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Rajeeva Karandikar*

Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi, India

Received 31 May 1993

Abstract: In the existing 'direct' white noise theory of nonlinear filtering, the state process is still modelled as a Markov process satisfying an Itô stochastic differential equation, while a 'finitely additive' white noise is used to model the observation noise. We remove this asymmetry by modelling the state process as the solution of a (stochastic) differential equation with a 'finitely additive' white noise as the input. This enables us to introduce correlation between the state and observation noises, and to obtain robust nonlinear filtering equations in the correlated noise case.

Keywords: Robust nonlinear filtering; white noise; cyclinder probability; Markov property.

1. Introduction

Nonlinear filtering equations are usually derived using the theory of stochastic differential equations as developed by Itô [7]. These equations hold, as is common in probability theory, almost surely in the space of continuous functions. The difficulty in applying this theory in practice is that the set of all possible real data has measure zero! In technical terms, nonlinear filtering equations obtained directly from Itô's theory are not robust. There are essentially two approaches to circumvent this difficulty. In one approach, due originally to Clark [4], the Itô equation of nonlinear filtering is rewritten in an equivalent pathwise form which does not involve differential of the observation process that appears in Itô's framework. This observation process is actually the integrated version of real data. The nonlinear filtering equation is then shown to be continuous (in some topology) with respect to the observation and, therefore, it can be extended by continuity to the actual sample paths [5]. In the other approach, due originally to Balakrishnan [2], one tries to model the observation process directly with a white noise error term. Although modelling white noise directly is intuitively appealing, it brings a host of mathematical complications. Kallianpur and Karandikar [9] developed the theory of nonlinear filtering in this framework. The advantage of this approach is that, once the mathematical difficulties are resolved, one always obtains results already in the robust form.

One drawback of the existing 'direct' white noise theory is that the state process in this theory is still modelled as a Markov process satisfying an Itô stochastic differential equation. The state and observation noises are then jointly modelled in an appropriate product space which makes them necessarily uncorrelated.

Correspondence to: Prof. A. Bagchi, Department of Applied Mathematics, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands.

* This work was performed when the author was visiting the University of Twente with financial support from the Dutch Network on Systems and Control.

SSDI 0167-6911(93)E0119-2

One purpose of this paper is to remove this asymmetry in the theory by modelling the state process as the solution of a stochastic differential equation with a 'directly modelled' white noise as the input. We show that the solution of such a stochastic differential equation is, indeed, a Markov process. With this formulation, it is straightforward to introduce correlation between the state and observation noises. The other purpose of the paper is to obtain robust nonlinear filtering equations for stochastic dynamical systems with correlated state and observation noises.

2. Mathematical preliminaries

Let H be a real separable Hilbert space, with inner product denoted by $(\cdot,\cdot)$ and norm by $\|\cdot\|$. Let $\mathscr{P}$ be the set of all orthogonal projections on H with finite-dimensional range. We introduce a partial ordering on $\mathscr{P}$ by defining $P_1 \leq P_2$ if (Range P_1) $\subseteq$ (Range P_2), for $P_1, P_2 \in \mathscr{P}$.

For $P \in \mathscr{P}$, let

$$\mathscr{C}_p = \{P^{-1}B \mid B \text{ Borel in (Range } P)\}$$

and define

$$\mathscr{C} \triangleq \bigcup_{P \in \mathscr{P}} \mathscr{C}_P .$$

Sets in $\mathscr{C}$ are called *cylinder sets* in H. Each $\mathscr{C}_P$, for fixed $P \in \mathscr{P}$, is a σ-algebra, but $\mathscr{C}$ is only an algebra.

A *cylinder probability n* on H is a finitely additive probability measure on $(H, \mathscr{C})$ such that, for all $P \in \mathscr{P}$, the restriction n_P of n to $\mathscr{C}_P$ is countably additive. The *canonical Gauss measure m* on $(H, \mathscr{C})$ is a cylinder probability on H such that, for all $h \in H$,

$$\int_H \exp\{\mathrm{i}(h, h_1)\}\, \mathrm{d}m(h_1) = \exp\{-\tfrac{1}{2}\|h\|^2\}. \tag{1}$$

The identity map $I:(H, \mathscr{C}, m) \to (H, \mathscr{C})$ is called a Gaussian white noise on H.

A function $f: H \to \mathbb{R}$ belongs to $\mathscr{L}(H, \mathscr{C}, n)$ if f is Borel-measurable and the net $\{f_P \mid P \in \mathscr{P}\}$ of functions defined by

$$f_P(h) = f(Ph) \tag{2}$$

is Cauchy in probability. Elements of $\mathscr{L}(H, \mathscr{C}, n)$ are called *n-accessible random variables*. It is easy to extend the definition to the case when $f: H \to \mathbb{R}^d$, or even when $f: H \to S$, where S is a complete separable metric space. We denote the corresponding classes of random variables by $\mathscr{L}(H, \mathscr{C}, n; \mathbb{R}^d)$ or $\mathscr{L}(H, \mathscr{C}, n; S)$, respectively. For $f \in \mathscr{L}(H, \mathscr{C}, n; S)$, it can be shown that the net of measures $\{n \circ f_P^{-1}, P \in \mathscr{P}\}$ converges to a countably additive measure on S. The limit is called the measure induced by f under n, and is denoted by $n \circ f^{-1}$.

By the class $\mathscr{L}^1(H, \mathscr{C}, n)$ of *integrable accessible random variables* we mean all elements $f \in \mathscr{L}(H, \mathscr{C}, n)$ for which

$$\int_{\mathbb{R}} |x|\, \mathrm{d}(n \circ f^{-1})(x) < \infty .$$

For $f \in \mathscr{L}^1(H, \mathscr{C}, n)$, we define

$$\int_H f \mathrm{d}n \triangleq \int_{\mathbb{R}} x\, \mathrm{d}(n \circ f^{-1})(x). \tag{3}$$

Another useful notion is the representation of a cylinder measure and the corresponding lifting into the representation space. In fact, this is how Gross [8] defined a *weak distribution* and showed that it is equivalent to a cylinder measure. A *representation* (L, Π_0) of the cylinder measure n is a mapping L from H into the class of all random variables on some countably additive probability space $(\Omega_0, \mathscr{A}_0, \Pi_0)$ such that

$$L(\alpha_1 h_1 + \alpha_2 h_2) = \alpha_1 L(h_1) + \alpha_2 L(h_2) \text{ a.s. } (\Pi_0)$$

and

$$n(h \mid (h, h_1) \le x) = \Pi_0(L(h_1) \le x)$$

for all $x \in \mathbb{R}$. A map $f: H \to S$ is called a *cylinder function* if f may be expressed as

$$f(h) = f_0((h, e_1), \ldots, (h, e_k))$$

for $e_1, \ldots, e_k \in H$ and $f_0: \mathbb{R}^k \to S$. For any such cylinder function f, define

$$R_n(f) \triangleq f_0(L(e_1), \ldots, L(e_k)).$$

Suppose now that $f \in \mathscr{L}(H, \mathscr{C}, n; S)$. Then the net $R_n(f_P)$, f_P a cylinder function defined by (2), can be shown to converge in probability. The limit is called the *n-lifting* of f and is denoted by $R_n(f)$. One can prove that

$$n \circ f^{-1} = \Pi_0 \circ [R_n(f)]^{-1}$$

and as a consequence, for $f \in \mathscr{L}^1(H, \mathscr{C}, n)$,

$$\int_H f \, \mathrm{d}n \equiv \int_{\Omega_0} R_n(f) \, \mathrm{d}\Pi_0. \tag{4}$$

In the white noise theory, we need to introduce appropriately the notion of conditional expectation, as we cannot work anymore with countably additive probability measures and with σ-algebras. Let n be a cylinder probability measure on $(H, \mathscr{C})$ and let H_1 be another real separable Hilbert space. A mapping $\phi: H \to H_1$ is called a *cylindrical mapping* if, for all $h_1 \in H_1$,

$$(\phi(\cdot), h_1)_1 \in \mathscr{L}(H, \mathscr{C}, n),$$

where $(\cdot, \cdot)_1$ is the inner product in H_1.

Now let (L, Π_0) be a representation of the cylinder measure n. Define L_1 on H_1 by

$$L_1(h_1) = R_n((\phi, h_1)_1). \tag{5}$$

Then for any cylinder set $C_1 \in \mathscr{C}_1$ of the form

$$C_1 = \{k_1 \in H_1 \mid ((k_1, h_1)_1, \ldots, (k_1, h_d)) \in B\},$$

$h_1, \ldots, h_d \in H_1$ and B a Borel set in $\mathbb{R}^d$, define

$$n_1(C_1) = \Pi_0((L_1(h_1), \ldots, L_1(h_d)) \in B);$$

n_1 thus defined is a cylinder measure on $(H_1, \mathscr{C}_1)$ and is defined to be the *measure induced by* ϕ *under* n, denoted by $n \circ \phi^{-1}$. (L_1, Π_0) is a representation of $n \circ \phi^{-1}$ and is called the representation induced by ϕ. Let R_{n_1} be the corresponding lifting of n_1 and define

$$\mathscr{U}(\phi) \equiv \mathscr{U}(H, \mathscr{C}, n; \phi) \triangleq \{g \in \mathscr{L}(H_1, \mathscr{C}_1, n_1) \mid R_{n_1}(g) = R_n(g \circ \phi)\}. \tag{6}$$

If $g \in \mathscr{U}(\phi)$, then

$$n_1 \circ g^{-1} = n \circ (g \circ \phi)^{-1}.$$

We would like to mention that $\mathscr{U}(\phi)$ has to be introduced simply because we are unable to prove that (6) holds for all $g \in \mathscr{L}(H_1, \mathscr{C}_1, n_1)$. Now we are in a position to define conditional expectation. Let $\phi: H \to H_1$ be a cylindrical mapping. For $f \in \mathscr{L}(H, \mathscr{C}, n)$, if there exists $g \in \mathscr{U}(\phi)$ such that

$$\int_H f 1_{C_1}(\phi) \, \mathrm{d}n = \int_H g(\phi) 1_{C_1}(\phi) \, \mathrm{d}n$$

for all $C_1 \in \mathscr{C}_1$, then g is defined to be the *conditional expectation* of f given ϕ, denoted by $E_n[f \mid \phi]$. Details may be found in [10].

From now on we shall only deal with the Hilbert space

$$H = L_2([0, T], \mathbb{R}^d)$$

and the canonical Gauss measure m on H. In our subsequent development we need the following result, which follows as a consequence of Theorem V.3.6. and the discussion in Section III. 4 of [10].

Theorem 1. *Let* $B \triangleq C_0([0, T], \mathbb{R}^d)$ *be the class of continuous functions on* $[0, T]$ *with the value at* $t = 0$ *being* 0, *and let* $\gamma : H \to B$ *be defined by*

$$\gamma(\eta)(t) = \int_0^t \eta(s)\,\mathrm{d}s, \quad \eta \in H. \tag{7}$$

If $f^{\sim}$ *is a real-valued continuous function on* B *(where* B *is equipped with the topology of uniform convergence), and* $f \triangleq f^{\sim} \circ \gamma$, *then*

$$f \in \mathscr{L}(H, \mathscr{C}, m).$$

Furthermore, if $\{W_t, t \geq 0\}$ *is an* $\mathbb{R}^d$*-valued standard Wiener process on some probability space* $(\Omega_0, \mathscr{A}_0, \Pi_0)$, *then* L *defined by*

$$L(\eta) = \int_0^T [\eta(s), \mathrm{d}W_s],$$

the vector Wiener integral, is a representation of m *and, if* R_m *is the corresponding m-lifting,*

$$R_m(f) = f^{\sim}(W.).$$

Moreover, if μ *is the Wiener measure on* B, $f \in \mathscr{L}^1(H, \mathscr{C}, m)$ *iff* $f^{\sim} \in \mathscr{L}^1(B, \mathscr{B}(B), \mu)$, $\mathscr{B}(B)$ *the Borel* σ*-algebra of* B, *and*

$$\int_H f\,\mathrm{d}m = \int_B f^{\sim}\,\mathrm{d}\mu.$$

3. A version of the Fubini theorem for the Gauss measure

A major result in the theory of countably additive measures is the Fubini theorem. In this section we obtain a version of the Fubini theorem for (finitely additive) Gauss measure which will be needed later. The result of this section is, moreover, of independent interest.

We begin with some notations. Let H, H_1, H_2 denote the Hilbert spaces

$$H = L_2([0, T], \mathbb{R}^d), \qquad H_1 = L_2([0, s], \mathbb{R}^d), \qquad H_2 = L_2([s, T], \mathbb{R}^d)$$

and let B, B_1, B_2 denote the Banach spaces

$$B = C_0([0, T]; \mathbb{R}^d), \qquad B_1 = C_0([0, s]; \mathbb{R}^d), \qquad B_2 = C_0([s, T], \mathbb{R}^d).$$

We define $\gamma : H \to B$ by

$$\gamma(\eta)(t) = \int_0^t \eta(\sigma)\,\mathrm{d}\sigma$$

and define $\gamma_1 : H_1 \to B_1$, $\gamma_2 : H_2 \to B_2$ by

$$\gamma_1(\eta_1)(t) = \int_0^t \eta_1(\sigma)\,\mathrm{d}\sigma, \qquad \gamma_2(\eta_2)(t) = \int_s^t \eta_2(\sigma)\,\mathrm{d}\sigma.$$

Let us define $Q_1: H \to H_1$, $Q_2: H \to H_2$, $Q_1^{\sim}: B \to B_1$, $Q_2^{\sim}: B \to B_2$ by

$$Q_1(\eta)(t) = \eta(t), \quad 0 \le t \le s; \; Q_2(\eta)(t) = \eta(t), \; s \le t \le T,$$

$$Q_1^{\sim}(\zeta)(t) = \zeta(t), \quad 0 \le t \le s; \; Q_2^{\sim}(\zeta)(t) = \zeta(t) - \zeta(s), \; s \le t \le T.$$

Finally, define $Q: H_1 \times H_2 \to H$, $Q^{\sim}: B_1 \times B_2 \to B$ by

$$Q(\eta_1, \eta_2)(t) = \begin{cases} \eta_1(t), & t \le s, \\ \eta_2(t), & t > s, \end{cases} \qquad Q^{\sim}(\zeta_1, \zeta_2)(t) = \begin{cases} \zeta_1(t), & t \le s, \\ \zeta_1(s) + \zeta_2(t), & t > s. \end{cases}$$

It is easy to see that Q_1, Q_2 are cylindrical mappings and that $m_1 = m \circ Q_1^{-1}$, $m_2 = m \circ Q_2^{-1}$ are the Gauss measures on H_1, H_2, respectively. Intuitively, one expects m to be the 'product measure' of m_1 and m_2 and that, for $f \in \mathscr{L}^1(H, \mathscr{C}, m)$.

$$\int f(\eta)\, \mathrm{d}m(\eta) = \int\int f(Q(\eta_1, \eta_2))\, \mathrm{d}m_2(\eta_2)\, \mathrm{d}m_1(\eta_1). \tag{8}$$

It is not clear as to whether (8) holds for all $f \in \mathscr{L}^1(H, \mathscr{C}, m)$. In [10] the validity of (3) has been proved for a certain subclass of $\mathscr{L}^1(H, \mathscr{C}, m)$. In the sequel we need the Fubini theorem for a different subclass. We begin with an auxiliary result.

Lemma 2. *Let $f_1: H_1 \to \mathbb{R}$ be a function satisfying*

$$f_1(\eta_1) = f_1^{\sim}(\gamma_1(\eta_1))$$

for some continuous function $f_1^{\sim}: B_1 \to \mathbb{R}$. Then

$$f_1 \in \mathscr{U}(H, \mathscr{C}, m; Q_1).$$

Proof. Let $\{W_t,\ 0 \le t \le T\}$ be an $\mathbb{R}^d$-valued standard Wiener process on some probability space (Ω_0, A_0, Π_0). Let us define

$$L(\eta) \triangleq \int_0^T [\eta(s), \mathrm{d}W_s], \quad \eta \in H.$$

Then (L, Π_0) is a representation of m. Since $f_1 \circ Q_1 = f_1^{\sim} \circ Q_1^{\sim} \circ \gamma$, and $f_1^{\sim} \circ Q_1^{\sim}$ is continuous function, it follows from Theorem 1 that

$$R_m(f_1 \circ Q_1) = f_1^{\sim}(Q_1^{\sim}(W_{\cdot})) = f_1^{\sim}(W'_{\cdot}),$$

where $W'_t = W_t$, $0 \le t \le s$.

On the other hand, the representation (L_1, Π_0) of m_1 induced by Q_1 is given by

$$L_1(\eta_1) = R_m((Q_1(\cdot), \eta_1)) = \int_0^s [\eta_1(s), \mathrm{d}W_s] = \int_0^s [\eta_1(s), \mathrm{d}W'_s]$$

and hence, again by Theorem 1,

$$R_{m_1}(f_1) = f_1^{\sim}(W'_{\cdot}).$$

Thus, $R_{m_1}(f_1) = R_m(f_1 \circ Q_1)$, implying that

$$f_1 \in \mathscr{U}(Q_1; H, \mathscr{C}, m). \qquad \square$$

This leads to the following version of Fubini's theorem.

Theorem 3. *Let $f: H \to \mathbb{R}$ satisfy $f = f^{\sim} \circ \gamma$ for some real-valued bounded continuous function $f^{\sim}$ on B. Then equality* (8) *holds. Furthermore, if*

$$f_1(\eta_1) \triangleq \int f(Q(\eta_1, \eta_2)) \, \mathrm{d}m_2(\eta_2) \tag{9}$$

then

$$E_m[f \mid Q_1] = f_1 \circ Q_1. \tag{10}$$

Proof. Using notations of the previous lemma, one has

$$R_m(f) = f^{\sim}(W.).$$

Define $f_1^{\sim}: B_1 \to \mathbb{R}$ by

$$f_1^{\sim}(\zeta_1) = E[f^{\sim}(Q^{\sim}(\zeta_1, W''))],$$

where $W'' = Q_2^{\sim} W$. Using the fact that, for fixed $\zeta_1, \zeta_2 \to f^{\sim}(Q^{\sim}(\zeta_1, \zeta_2))$ is continuous, it follows from Theorem 1 that $f_1(\eta_1) = f^{\sim}(\gamma_1(\eta_1))$. This implies, again by Theorem 1, that

$$R_{m_1}(f_1) = f_1^{\sim}(W'.).$$

It is easy to verify that $f_1^{\sim}$ is a continuous function. This fact, with Lemma 2, implies that

$$R_m(f_1 \circ Q_1) = R_{m_1}(f_1) = f_1^{\sim}(W'.). \tag{11}$$

Note that $f^{\sim}(Q^{\sim}(W', W'')) = f^{\sim}(W)$. Clearly,

$$E[f^{\sim}(W) \mid W_t, 0 \le t \le s] = f_1^{\sim}(W'). \tag{12}$$

Equalities (10), (11) and Theorem IV. 3.2 in [10] yield

$$E_m[f \mid Q_1] = f_1 \circ Q_1.$$

Also, we have here $Ef_1^{\sim}(W') = Ef^{\sim}(W)$ and hence

$$\int f_1(\eta_1) \, \mathrm{d}m_1(\eta_1) = \int f(\eta) \, \mathrm{d}m(\eta).$$

This establishes equality (3). □

4. Markov property

Let $u: \mathbb{R}^d \to \mathbb{R}^d$ be a Lipschitz function. For $x_0 \in \mathbb{R}^d$ and $\eta \in H$, consider the following differential equation:

$$\frac{\mathrm{d}x_t(\eta)}{\mathrm{d}t} = u(x_t(\eta)) + e_t(\eta), \tag{13}$$

where e defined by $e_t(\eta) = \eta(t)$ is a white noise on H. We shall presently see that equation (13) admits a unique solution. We shall first show that (13) is equivalent to

$$x_t(\eta) = x_0 + \int_0^t u(x_\tau(\eta)) \, \mathrm{d}\tau + \int_0^t e_\tau(\eta) \, \mathrm{d}\tau. \tag{14}$$

More generally, consider for $x_0 \in \mathbb{R}^d$ and $\zeta \in B = C([0, T], \mathbb{R}^d)$,

$$X_t = x_0 + \int_0^t u(X_\tau) \, \mathrm{d}\tau + \zeta(t). \tag{15}$$

Existence and uniqueness of the solution to this equation follows from the Lipschitz condition on u. In fact, uniqueness may be established easily. If $\{X_t'\}$ is another solution of (15) with $\zeta' \in B$ as the forcing term, i.e.

$$X_t' = x_0 + \int_0^t u(X_\tau')\,\mathrm{d}\tau + \zeta'(t),$$

then

$$\sup_{\tau \le t} |X_\tau - X_\tau'| \le K \int_0^t |X_\tau - X_\tau'|\,\mathrm{d}\tau + \sup_{\tau \le t} |\zeta(\tau) - \zeta'(\tau).$$

Using Gronwall's inequality, we have

$$\sup_{\tau \le t} |X_\tau - X_\tau'| \le K_1 \sup_{\tau \le t} |\zeta(\tau) - \zeta'(\tau)|. \tag{16}$$

Uniqueness of the solution of (15) readily follows from this. Denoting the solution of (15) by $X_t(\zeta)$, estimate (16) implies that $\zeta \to X_t(\zeta)$ is continuous from B into $\mathbb{R}$ and $\zeta \to X_\cdot(\zeta)$ is continuous from B into B. Moreover,

$$x_t(\eta) = X_t(\gamma(\eta)), \tag{17}$$

with γ defined in (7). Therefore, from Theorem 1,

$$x_t \in \mathscr{L}(H, \mathscr{C}, m) \quad \text{and} \quad x \in \mathscr{L}(H, \mathscr{C}, m; B).$$

We are now in a position to establish the Markov property.

Theorem 4. *Let g be a bounded continuous function on $\mathbb{R}^d$. Then, for $t \ge s$,*

$$E_m[g(x_t)|Q_1] = g_1(x_s), \tag{18}$$

where

$$g_1(x) = \int g(\Gamma_{st}(x, \eta_2))\,\mathrm{d}m_2(\eta_2), \tag{19}$$

with $x_t' \triangleq \Gamma_{st}(x, \eta_2)$ being the unique solution of

$$x_t' = x + \int_0^t u(x_\tau')\,\mathrm{d}\tau + \int_s^t e_\tau(\eta)\,\mathrm{d}\tau, \quad t \ge s. \tag{20}$$

Proof. Let $f(\eta) \triangleq g(x_t(\eta))$ and $f^\sim(\zeta) \triangleq g(X_t(\zeta))$. Then $f^\sim$ is a continuous function and $f = f^\sim \circ \gamma$. Therefore, by Theorem 3

$$E_m[f|Q_1] = f_1 \circ Q_1$$

exists. We now show that

$$f_1 \circ Q_1 = g_1(x_1). \tag{21}$$

For $\eta_1 \in H_1$, let $\xi_\tau(\eta_1)$, $\tau \le s$, be the solution of

$$\xi_\tau = x_0 + \int_0^\tau u(\xi_\sigma)\,\mathrm{d}\sigma + \int_0^\tau \eta_1(\sigma)\,\mathrm{d}\sigma.$$

Then

$$x_s(\eta) = \xi_s(Q_1\eta).$$

Also

$$x_t(Q(\eta_1, \eta_2)) = \Gamma_{st}(\xi_s(\eta_1), \eta_2).$$

Therefore,

$$f_1(\eta_1) = \int g(\Gamma_{st}(\xi_s(\eta_1), \eta_2))\,dm_2(\eta_2) = g_1(\xi_s(\eta_1)).$$

Then

$$f_1(Q_1\eta) = g_1(\xi_s(Q_1\eta)) = g_1(x_s(\eta)),$$

proving the Markov property (18). □

Remark. In the usual countably additive set-up, Theorem 4 would imply that

$$E[g(x_t)|\mathscr{F}_s^x] = E[g(x_t)|x_s],$$

where g is a bounded continuous function on $\mathbb{R}^d$ and $\mathscr{F}_s^x$ is the smallest σ-algebra generated by $x_\sigma, 0 \le \sigma \le s$. It is well-known that this equality is equivalent to the usual definition of the process $\{x_t\}$ being Markov.

5. Robust filtering with correlated state and observation noises

We now consider the filtering problem where the state is the solution of the stochastic differential equation studied in Section 4 and the observation is corrupted by another white noise, possibly correlated with the state noise. We study only the scalar case of $d = 1$.

We take $H = L_2([0, T]; \mathbb{R}^2]$ and, for $\eta \in H$ expressed as $\eta(t) = (\eta_1(t), \eta_2(t))$, we define

$$e_{1t}(\eta) \triangleq \eta_1(t),$$
$$e_{2t}(\eta) \triangleq \eta_2(t).$$

We consider the following filtering model. The signal process $\{x_t(\eta)\}$ is given by

$$\begin{aligned} \dot{x}_t(\eta) &= u(x_t(\eta)) + \alpha e_{1t}(\eta) + \delta e_{2t}(\eta), \\ x_0(\eta) &= x_0, \end{aligned} \tag{22}$$

where $u: \mathbb{R} \to \mathbb{R}$ is a bounded Lipschitz function and $\alpha > 0$, $\delta > 0$ are real numbers. The observation process $\{y_t(\eta)\}$ is given by

$$y_t(\eta) = h(x_t(\eta)) + e_{1t}(\eta), \tag{23}$$

where h is assumed to be a twice continuously differentiable function on $\mathbb{R}$ such that h and h' are bounded. As noted in the previous section, the solution process $\{x_t\}$ of (13) is a Markov process on $(H, \mathscr{C}, m)$, m being the canonical Gauss measure on H. The nonlinear filtering problem consists of obtaining a formula for

$$\pi_t(f, y) \triangleq E_m[f(x_t)|Q_t y], \tag{24}$$

where $Q_t: H \to H$ is defined by $(Q_t\eta)(s) = \eta(s)1_{[0,t]}(s)$. This is intuitively clear as the right-hand side of (24) is the 'white noise' analog of the expression $E(f(x_t)|y_s; 0 \le s \le t]$ calculated in the usual set-up. Taking a cue from the case of independent signal and noise in the white noise approach adopted here [10], as well as standard results on nonlinear filtering in the usual 'countably additive' approach, one expects the following result. Let

$$\rho_t(\eta) \triangleq \exp\left\{\int_0^t h(x_s(\eta))y_s(\eta)\,ds - \frac{1}{2}\alpha\int_0^t h'(x_s(\eta))\,ds - \frac{1}{2}\int_0^t h^2(x_s(\eta))\,ds\right\}. \tag{25}$$

Then

$$\pi_t(f, y) = \sigma_t(f, y)/\sigma_t(1, y), \tag{26}$$

where

$$\sigma_t(f, y) \triangleq E_0[f(x_t)\rho_t | Q_t y], \tag{27}$$

where $E_0(\cdot | Q_t y)$ is the conditional expectation under the measure m_0, defined by

$$m_0(C) = \int_C \rho_T^{-1}\, dm, \quad C \in \mathscr{C}. \tag{28}$$

To evaluate the desired conditional expectation, we follow the approach of Davis [6] to pathwise filtering (Stratonovich form) in the 'countably additive' setting. It is not directly clear whether this form of the filtering is also robust, i.e. continuous with respect to the observation path. We show robustness directly in the white noise framework. For an indirect proof of robustness in the 'countably additive' setting, see [3].

We introduce a new process $\{z_t(\eta)\}$ by

$$\dot{z}_t(\eta) = \dot{x}_t(\eta) - \alpha y_t(\eta),$$
$$z_0(\eta) = x_0. \tag{29}$$

Then

$$z_t(\eta) = v_t(y(\eta), e_2(\eta)), \tag{30}$$

where, for $\phi_1, \phi_2 \in \bar{H} = L_2([0, T] : \mathbb{R})$, $v_t \equiv v_t(\phi_1, \phi_2)$ is the unique solution of

$$v_t = x_0 + \int_0^t [u(v_s + \alpha\hat{\phi}_1(s)) - \alpha h(v_s + \alpha\hat{\phi}_1(s)]\, ds + \delta \int_0^t \phi_2(s)\, ds, \tag{31}$$

with $\hat{\phi}_1(s) \triangleq \int_0^s \phi_1(\tau)\, d\tau$. Let $\bar{m}$ be the canonical Gauss measure on $\bar{H}$. Below follows our main result on robust nonlinear filtering for the case of correlated state and observation noises.

Theorem 5. *For $\bar{\phi} \in \bar{H}$, define*

$$\sigma_t(f, \bar{\phi}) = \int_{\bar{H}} f(v_t(\phi_1, \phi_2) + \alpha\hat{\phi}_1(t)) \exp\left\{ \int_0^t h(v_s(\phi_1, \phi_2) + \alpha\hat{\phi}_1(s))\phi_1(s)\, ds \right.$$
$$\left. - \frac{1}{2}\int_0^t (\alpha h' + h^2) v_s(\phi_1, \phi_2) + \alpha\hat{\phi}_1(s)\, ds \right\} d\bar{m}(\bar{\phi}). \tag{32}$$

Then we have

$$E_m[f(x_t) | Q_t y] = \frac{\sigma_t(f, y)}{\sigma_t(1, y)}. \tag{33}$$

Proof. One can prove this result following the procedure outlined above. It is, however, easier to follow a different approach which makes use of results on lifting. For this purpose, we need to introduce more notation. As in Section 2, we take $B = C_0([0, T]; \mathbb{R}^2)$ and, for $\eta \in H$,

$$\gamma(\eta)(t) = \int_0^t \eta(s)\, ds \tag{34}$$

so that $\gamma : H \to B$. Let μ be the Wiener measure on B. For $\zeta \in B$, $\zeta = (\zeta_1, \zeta_2)$, let

$$\beta_{1t}(\zeta) = \zeta_1(t), \qquad \beta_{2t}(\zeta) = \zeta_2(t)$$

so that $\{\beta_{1t}\}, \{\beta_{2t}\}$ are independent Wiener processes on B. For $\zeta \in B$, let $X_t(\zeta)$ be the unique solution of

$$X_t(\zeta) = x_0 + \int_0^t u(X_s(\zeta))\, ds + \alpha\beta_{1t}(\zeta) + \delta\beta_{2t}(\zeta). \tag{35}$$

As noted in Section 4, $X_t(\zeta)$ is a continuous function of ζ and $x_t(\eta) \equiv X_t(\gamma(\eta))$. Thus

$$R_m(x_t) = X_t. \tag{36}$$

Let us define

$$Y_t(\zeta) = \int_0^t h(X_s(\zeta))\,\mathrm{d}s + \beta_{1t}(\zeta). \tag{37}$$

Then the continuity of $X_s(\zeta)$ in ζ implies that $Y_t(\zeta)$ is also a continuous function of ζ and, moreover,

$$\int_0^t y_s(\eta)\,\mathrm{d}s = Y_t(\gamma(\eta)). \tag{38}$$

Hence we have

$$R_m\left(\int_0^t y_s\,\mathrm{d}s\right) = Y_t. \tag{39}$$

Our next step is to invoke part (iv) of Theorem IV-4.5 in [10]. If $\tilde{g}: B \to \mathbb{R}$ is a continuous function with the property that

$$\tilde{g}(\zeta) = \tilde{g}(\zeta') \text{ whenever } \zeta(s) = \zeta'(s) \text{ for all } s \le t,$$

then it follows that $g(n) \triangleq \tilde{g}(\gamma(\eta))$ satisfies $g(Q_t\eta) = g(\eta)$. Using Theorem 1, it can be proved that

$$g \in \mathscr{U}(Q_t y).$$

Also, the σ-algebra $\bar{\mathscr{D}}_{Q_t y}$ in [10, Theorem IV-4.5] can be identified as

$$\bar{\mathscr{D}}_{Q_t y} \equiv \sigma(Y_s;\ 0 \le s \le t),$$

where the right-hand side means the smallest σ-algebra generated by $\{Y_s, 0 \le s \le t\}$. Thus, if we could obtain a continuous function $\tilde{g}: B \to \mathbb{R}$ such that

$$\hat{\pi}_t(f, Y) = E_\mu[f(X_t) | \sigma(Y_s;\ 0 \le s \le t)] = \tilde{g}(Y), \tag{40}$$

then it follows from [10, Theorem IV-4.5] that

$$\pi_t(f, y) = \tilde{g}(\gamma(y)). \tag{41}$$

We, therefore, concentrate on $\hat{\pi}_t(f, Y)$. Standard results on nonlinear filtering yield

$$\hat{\pi}_t(f, Y) = \hat{\sigma}_t(f, Y)/\hat{\sigma}_t(1, Y), \tag{42}$$

where

$$\hat{\sigma}_t(f, Y) = E_0[f(X_t)\hat{\rho}_t | \sigma(Y_s;\ 0 \le s \le t)], \tag{43}$$

$$\hat{\rho}_t = \exp\left\{\int_0^t h(X_s)\,\mathrm{d}Y_s - \frac{1}{2}\int_0^t h^2(X_s)\,\mathrm{d}s\right\} \tag{44}$$

and E_0 is the expectation operation with respect to the measure P_0 defined by

$$\frac{\mathrm{d}P_0}{\mathrm{d}P} = \rho_T^{-1}. \tag{45}$$

Also, under P_0, $\{Y_t\}$ and $\{\beta_{2t}\}$ are independent processes and $\{\beta_{2t}\}$ continues to be a Wiener process under P_0.

Now, for $\phi_1, \phi_2 \in C([0,T], \mathbb{R})$, define $V_t = v_t(\phi_1, \phi_2)$ as the solution of

$$V_t = x_0 + \int_0^t [u(V_s + \alpha\phi_1(s)) - \alpha h(V_s + \alpha\phi_1(s))]\,\mathrm{d}s + \delta\phi_2(s), \tag{46}$$

and let

$$Z_t \triangleq V_t(Y_\cdot, \beta_2). \tag{47}$$

It is easily seen that $V_t(\phi_1, \phi_2)$ is a continuous function of ϕ_1, ϕ_2. Then we have the pathwise formula of nonlinear filtering [6]

$$\begin{aligned}\hat{\sigma}_t(f, Y) = \int_{\bar{B}} f(Z_t + \alpha Y_t) \exp\Bigg\{ \int_0^{Y_t} h(Z_t + \alpha s)\,\mathrm{d}s - \int_0^t (\alpha h' + h^2)(Z_s + \alpha Y_s)\,\mathrm{d}s \\ - \frac{1}{2}\int_0^t [h(Z_s + \alpha Y_s) - h(Z_s)][u(Z_s + \alpha Y_s) - \alpha h(Z_s + \alpha Y_s)]\,\mathrm{d}s \\ - \frac{1}{2}\int_0^t [h(Z_s + \alpha Y_s - h(Z_s)]\,\mathrm{d}\beta_{2s}\Bigg\}\,\mathrm{d}\bar{\mu},\end{aligned} \tag{48}$$

where $\bar{\mu}$ is the Wiener measure on $\bar{B} = C([0,T], \mathbb{R})$. Recall that $Z_t = V_t(Y, \beta_2)$ and Y is treated as a parameter in the integral above. It follows that the functionals $\sigma_t(f, \bar{\phi})$, defined via (32), and $\hat{\sigma}_t(f, Y)$, defined via (43), are related by

$$\sigma_t(f, y) \equiv \hat{\sigma}_t(f, \gamma(y)). \tag{49}$$

To complete the proof of Theorem 5, it suffices to prove that the pathwise filter $\hat{\sigma}_t(f, Y)$ given by (48) is a continuous function of Y.

Using the assumptions that u, h, h' are bounded continuous functions, it follows that, for $Y^n \to Y$ in $C([0,T], \mathbb{R})$,

$$V(Y^n, \beta_2) \to V(Y, \beta_2) \quad \text{in } \bar{\mu}\text{-probability}.$$

Since h is assumed to be bounded, we may use standard results on stochastic integrals to show that

$$\begin{aligned}&\int_0^t [h(V_s(Y^n, \beta_2) + \alpha Y_s^n) - h(V_s(Y^n, \beta_2))]\,\mathrm{d}s \\ &\quad \to \int_0^t [h(V_s(Y, \beta_2) + \alpha Y_s) - h(V_s(Y, \beta_2))]\,\mathrm{d}s \quad \text{in } \bar{\mu}\text{-probability}.\end{aligned}$$

Boundedness of u, h, h' now give

$$\hat{\sigma}_t(f, Y^n) \to \hat{\sigma}_t(f, Y),$$

completing the proof of Theorem 5. □

6. Conclusion

We proved two important results on the direct modelling of white noise in stochastic systems. We showed that if we model the state process directly with the 'finitely additive' white noise as the forcing term, the solution of the resulting stochastic differential equation is a Markov process. This led to the modelling of a stochastic dynamical system with correlated state and observation noises. We then obtained the robust nonlinear filtering formula for this correlated case. It may be interesting to study Radon–Nikodym derivatives for the case of correlated state and observation noises. The uncorrelated case has been studied in [1].

Acknowledgment

The authors are grateful to Prof. J.M.C. Clark for pointing out the paper mentioned in [3]. The comments of an anonymous referee are also appreciated.

References

[1] A. Bagchi and R.R. Mazumdar, On Radon–Nikodym derivatives of finitely additive measures induced by nonlinear transformations on Hilbert space, to appear: *J. Nonlinear Anal. Theory Methods Applic.*

[2] A.V. Balakrishnan, Nonlinear white noise theory, in: P.R. Krishnaiah, ed., *Multivariate Analysis – V* (North-Holland, Amsterdam, 1980) 97–109.

[3] M. Chaleyat-Maurel and D. Michel, Une propriété de robustesse en filtrage nonlinéaire, *Stochastics* **19** (1986) 11–40.

[4] J.M.C. Clark, The design of robust approximations to the stochastic differential equations of nonlinear filtering, in: J.K. Skwirzynski, ed., *Communication Systems and Random Process Theory*, NATO Advanced Study Institute Series (Sijthoff and Noordhoff, Alphen aan den Rijn, 1978).

[5] M.H.A. Davis, On a multiplicative functional transformation arising in nonlinear filtering theory, *Z. Wahrsch. Verw. Geb.* **54** (1980) 125–139.

[6] M.H.A. Davis, Pathwise non-linear filtering, in: M. Hazewinkel and J.C. Willems, eds., *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, NATO Advanced Study Institute Series (Reidel, Dordrecht, 1980).

[7] M.H.A. Davis and S.I. Marcus, An introduction to nonlinear filtering, in: M. Hazewinkel and J.C. Willems, eds., *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, NATO Advanced Study Institute Series (Reidel, Dordrecht, 1980).

[8] L. Gross, Abstract Wiener space, in: *Proc. 5th Berkeley Symp. on Mathematical Statistics and Probability*, Berkeley, Vol. 2 (1965) 31–42.

[9] G. Kallianpur and R.L. Karandikar, A finitely additive white noise approach to nonlinear filtering, *Appl. Math. Optim.* **10** (1983) 159–186.

[10] G. Kallianpur and R.L. Karandikar, *White Noise Theory of Filtering, Prediction and Smoothing* (Gordon and Breach, London, 1988).